UNIVERSITY of TASMANIA

School of Natural Sciences

College of Science and Engineering

KMA712
Introduction to Bioinformatics

Semester 1, 2019

Unit Outline

A/Prof Michael Charleston

CRICOS Provider Code: 00586B

CONTACT DETAILS

Unit coordinator

Unit coordinator: A/Prof Michael Charleston

Campus: Sandy Bay

Email: michael.charleston@utas.edu.au

Phone: (03) 6226 2444

Room location and number: Maths & Physics Building, Clark Road, Rm443

Consultation hours: Open door policy

WHAT IS THE UNIT ABOUT?

Unit description

Bioinformatics is a core skill that is needed by practitioners in the life sciences, from health to pathology and epidemiology to fundamental biological research. It is a discipline that brings together tools from mathematics, statistics, and computer science, to answer challenging questions that arise in the modern life sciences.

The unit is aimed at those pursuing research in the life sciences and/or working with large biological data sets, and is intended to provide them with the practical skills and knowledge to effectively use existing bioinformatics tools, to understand how the tools work and what can or cannot be done with them, and to develop and critically assess bioinformatics workflows and methodologies.

Topics include relevant biological background, use of command-line tools and bioinformatics software, biological sequence comparison and alignment, maximum likelihood, phylogenetic inference, genome assembly and mapping, and RNA-Seq expression analysis. Some flexibility of topics is possible depending on class requirements.

Assignments may be based on students' own data, subject to unit coordinator approval.

These tools and skills will enhance students' research capacity in life sciences.

Intended Learning Outcomes

On successful completion of this unit, students will be able to

1. Effectively use a range of *current bioinformatics tools and workflows* to solve problems in life sciences ("software");
2. Explain and apply the fundamental *computational methods* underlying the solution of bioinformatics problems ("methods");
3. Classify the *major classes* of bioinformatics problems and choose appropriate tools to solve them ("major problems");
4. Evaluate and *communicate* bioinformatics methodologies clearly ("communication").

Graduate Quality Statement

Successful completion of this unit supports your development of course learning outcomes, which describe what a graduate of a course knows, understands and is able to do. Course learning outcomes are available from the Course Coordinator. Course learning outcomes are developed with reference to national discipline standards, Australian Qualifications Framework (AQF), any professional accreditation requirements and the University of Tasmania's Graduate Quality Statement.

> The University of Tasmania experience unlocks the potential of individuals. Our graduates are equipped and inspired to shape and respond to the opportunities and challenges of the future as accomplished communicators, highly regarded

professionals and culturally competent citizens in local, national, and global society. University of Tasmania graduates acquire subject and multidisciplinary knowledge and skills and develop creative and critical literacies and skills of inquiry. Our graduates recognise and critically evaluate issues of social responsibility, ethical conduct and sustainability. Through respect for diversity and by working in individual and collaborative ways, our graduates reflect the values of the University of Tasmania.

Alterations to the unit as a result of student feedback

This unit is a re-design of the previous version, which was a short intensive "block" course. In response to feedback, this redesign is more spread out to give students time to absorb material better and to give more time for practising the skills that are being taught.

Prior knowledge and/or skills

Assumed knowledge:

Students enrolling in it should have a BSc or equivalent.

Prior to beginning the unit you should have at least a *basic knowledge* of the fundamentals of molecular biology (such as that covered by KPZ164), in particular the structure of DNA and RNA, the fundamental dogma of molecular biology, and essential evolutionary theory. You should have an *appreciation* of simple mathematical terminology and symbols and be able to carry out simple statistical analyses. *If you are in doubt about whether you have the appropriate background you should contact the Unit Coordinator.*

It is *not* assumed that you will have any programming experience or advanced understanding in mathematics or statistics.

HOW WILL I BE ASSESSED?

Assessment schedule

Assessment task	Date due	Percent weighting	Intended Learning Outcomes measured
Assessment Task 1: Written quiz	April 2nd in lab	20%	ILO 2, 3, 4
Assessment Task 2: Practical exercises	Each week	15%	ILO 1, 2
Assessment Task 3: Assignment 1	April 26th 5pm	25%	ILO 1, 2, 4
Assessment Task 4: Assignment 2	June 14th 5pm	40%	ILO 1, 2, 3, 4

Assessment details

Assessment Task 1: Written quiz

Description / conditions	This pen-and-paper quiz will be for students to demonstrate that they understand, and can communicate, the concepts and methods that have been covered in the material up to the previous week. **Students must attend the quiz in person.**
Assessment criteria	Questions will be asked on the content covered up to the week before the quiz. A combination of modes of question will be used, e.g., true / false, short answer, multiple choice, and demonstrations of algorithms.
Intended Learning Outcomes measured	ILO2 (methods) ILO3 (major problems) ILO4 (communication)
Duration	45 minutes.
Date	Mid-semester

Assessment Task 2: Practical exercises

Task description	These computer lab-based exercises are designed to help students get hands-on experience in applying their knowledge to solve real problems in bioinformatics. They will work in a lab or remotely with online help from the Unit Coordinator or agent thereof, with time to complete the exercises before submission for assessment if required. **Students are required to attend the majority of these practicals in person.**
Assessment criteria	The assessment will be graded according to the following criteria: Correct use of software tools for the task, and demonstrated understanding of their relevance and function. Expertise will be judged by the lab demonstrator / unit coordinator for each lab and used to guide an overall grade for the semester.
Intended Learning Outcomes measured	ILO1 (software) ILO2 (methods)
Duration	During each lab class, up to 2 hours.
Date due	Each week.

Assessment Task 3: Assignment 1

Task description	Assignment 1 is to use a given set of bioinformatics tools to perform an analysis on a publicly available data set, for example, to find the set of all possible best matches between a set of query sequences and a set of target sequences, and estimate the phylogenetic relationships among the species found. The analysis will require students to use data from a public database, use current tools such as BLAST to identify sets of related sequences, to infer phylogenetic trees using a range of standard methods, and to interpret the overall results. The results should be presented clearly, critically assessed, and should be communicated with an intent that they be easily repeatable.
Assessment criteria	The assessment will be hand graded according to the following criteria: Software (ILO1): appropriate choice and correct use of software tools for the task. Methods (ILO2): quality of the way in which the analysis was carried out and interpretation of results Communication (ILO4): clarity, organization of the exposition and ease of repeatability.
Intended Learning Outcomes measured	ILO1 (software) ILO2 (methods) ILO4 (communication)
Task length	At least 5 pages including figures; around 1500 words
Date due	Approximately Week 8-9.

Assessment Task 4: Assignment 2

Task description	Assignment 2: The second assignment is to design and implement a repeatable bioinformatics workflow in order to answer a set of questions regarding the students' own data (or other data if required). Some of these questions will be chosen from the "major classes" of bioinformatics problems, being the most appropriate one(s) for the data set used. The methods used to solve the problems will be thoroughly explained as to a scientific, but not necessarily bioinformatics-focussed, audience. The assignment will describe the analysis and provide a critical discussion of the results. The workflow that is created must be shared in such a way that it can be run by others on the same data, e.g., using a public bioinformatics platform such as Galaxy.
Assessment criteria	The assessment will be graded according to the following criteria: Software (ILO1): design of workflow; Methods (ILO2): evidence of understanding of the methodologies used; Major Problems (ILO3): understanding and communication of this part of the analysis – dependent on the nature of the students' data;

	Communication (ILO4): clarity of the exposition and ease of repeatability
Intended Learning Outcomes measured	ILO1 (software) ILO2 (methods) ILO 3 (major problems) ILO4 (communication)
Task length	At least 8 pages including figures; around 2500 words
Date due	Examination period

How your final result is determined

In order to pass this unit you must:

1. gain an overall mark of at least 50%, and
2. demonstrate achievement of each of the intended learning outcomes 1-4. To demonstrate this you must (i) gain at least 45% in the combined mark from the assignments, and (ii) gain at least 45% in the combined mark from the quiz and the lab assessment.

In special circumstances ILOs may be assessed with alternative methods, with the agreement of the unit coordinator.

Submission of assignments

Submission of assignments will be on the MyLO website for this unit. Written documents must all be in .pdf format; if there are multiple files for a submission these must be compressed as .zip archives.

Requests for extensions

Requests for extensions should be made in writing to the Unit Coordinator; all reasonable requests will be considered.

Penalties

Work submitted late without permission for an extension will incur a late penalty of 10% reduction in the mark given per day or part thereof. For example an assignment submitted up to 24 hours (1 day) late that was otherwise worth 20 marks would be given a final mark of $20*0.9 = 18$ marks, and if it is submitted between 24 and 48 hours late it would get a final mark of $20*0.9^2 = 20*0.81 = 16.2 \rightarrow 16$. Work submitted more than 7 days late without permission for an extension will not receive a mark.

Review of results and appeals

If you wish to appeal a decision of the Unit Coordinator on a grade then you must make your appeal in writing to the Head of Discipline of Mathematics within 7 working days of receiving your grade.

Academic referencing

In your written work you will need to support your ideas by referring to scholarly literature, works of art and/or inventions. It is important that you understand how to correctly refer to the work of others, and how to maintain academic integrity.

Failure to appropriately acknowledge the ideas of others constitutes academic dishonesty (plagiarism), a matter considered by the University of Tasmania as a serious offence.

The appropriate referencing style for this unit is the "abbrv" BibTeX style, available from the Comprehensive TeX Archive Network (CTAN). This style has citations in brackets and numbered in the order in which they appear in the text; the bibliography has authors names displayed as initials-surname. If three or more authors are given just put the first one followed by "*et al.*". If submitting work using a word processor such as Word, the equivalent citation style should be used if possible.

The University library provides information on presentation of assignments, including referencing styles and should be referred to when completing tasks in this unit.

Please read the following statement on plagiarism. Should you require clarification please see your unit coordinator or lecturer.

Plagiarism

Plagiarism is a form of cheating. It is taking and using someone else's thoughts, writings or inventions and representing them as your own; for example, using an author's words without putting them in quotation marks and citing the source, using an author's ideas without proper acknowledgment and citation, copying another student's work.

If you have any doubts about how to refer to the work of others in your assignments, please consult your lecturer or tutor for relevant referencing guidelines. You may also find the Academic Honesty site on MyLO of assistance.

The intentional copying of someone else's work as one's own is a serious offence punishable by penalties that may range from a fine or deduction/cancellation of marks and, in the most serious of cases, to exclusion from a unit, a course or the University.

The University and any persons authorised by the University may submit your assessable works to a plagiarism checking service, to obtain a report on possible instances of plagiarism. Assessable works may also be included in a reference database. It is a condition of this arrangement that the original author's permission is required before a work within the database can be viewed.

For further information on this statement and general referencing guidelines, see the Plagiarism and Academic Integrity page on the University web site or the Academic Honesty site on MyLO.

Academic misconduct

Academic misconduct includes cheating, plagiarism, allowing another student to copy work for an assignment or an examination, and any other conduct by which a student:

a. seeks to gain, for themselves or for any other person, any academic advantage or advancement to which they or that other person are not entitled; or

b. improperly disadvantages any other student.

Students engaging in any form of academic misconduct may be dealt with under the Ordinance of Student Discipline, and this can include imposition of penalties that range from a deduction/cancellation of marks to exclusion from a unit or the University. Details of penalties that can be imposed are available in Ordnance 9: Student Discipline – Part 3 Academic Misconduct.

WHAT LEARNING OPPORTUNITIES ARE THERE?

MyLO

MyLO is the online learning environment at the University of Tasmania. This is the system that will host the online learning materials and activities for this unit.

Getting help with MyLO

It is important that you are able to access and use MyLO as part of your study in this unit. To find out more about the features and functions of MyLO, and to practice using them, visit the Getting Started in MyLO unit. For access to information about MyLO and a range of step-by-step guides in pdf, word and video format, visit the MyLO Student Support page on the University website.
If something is not working as it should, first check the current issues page in case this is a known problem; otherwise, contact the service desk (http://www.utas.edu.au/service-desk, e-mail service-desk@utas.edu.au, or phone 6226 1818), or request IT help online.

Resources

Required readings

There are no required readings for this course.

Recommended readings

Recommended readings will be provided as new publications appear in this rapidly changing field. In general, the journal Briefings in Bioinformatics is a good starting point for most topics covered.

Reading Lists

Reading Lists provide direct access to all material on unit reading lists in one place. This includes eReadings and items in Reserve. You can access the Reading List for this unit from the link in MyLO, or by going to on the University Library website.

Equipment, materials, software, accounts

All necessary equipment, software and accounts will be provided.

Activities

Learning expectations

The University is committed to high standards of professional conduct in all activities, and holds its commitment and responsibilities to its students as being of paramount importance. Likewise, it holds expectations about the responsibilities students have as they pursue their studies within the special environment the University offers.

The University's Code of Conduct for Teaching and Learning states:

Students are expected to participate actively and positively in the teaching/learning environment. They must attend classes when and as required, strive to maintain steady progress within the subject or unit framework, comply with workload expectations, and submit required work on time.

Details of teaching arrangements

The unit will be delivered face to face on campus in Sandy Bay. Lectures will be recorded. There are some assessments that require students to be present: the mid-semester quiz, and at least a majority of the lab exercises.

MyLO will be used as a focus for online discussions until the end of the course.

Specific attendance/performance requirements

Students must attend the mid-semester quiz and a majority of the lab sessions.

Teaching and learning strategies

The *face to face* delivery of this course in two-hour blocks for both lectures and labs, coupled with continuing support through online discussions and forums, is a compromise design to fit in with requirements of research scientists and practitioners who are unable to commit long periods to a study course, yet allowing students time to absorb the quite broad content of this unit.

Work Health and Safety (WHS)

The University is committed to providing a safe and secure teaching and learning environment. In addition to specific requirements of this unit you should refer to the University's and policy.

Communication

News and announcements will all be via MyLO; students will be expected to be aware of the content of such posts within 48 hours of them being posted.

Further information and assistance

If you are experiencing difficulties with your studies or assignments, have personal or life-planning issues, disability or illness which may affect your course of study, you are advised to raise these with the unit coordinator in the first instance.

There is a range of University-wide support services available to you including Student Learning Support, Student Advisers, Disability Services, and more which can be found on the Student Support and Development page of the University website.

Should you require assistance in accessing the Library, visit the Library website for more information.

Unit schedule

WEEK	TOPIC/ MODULE/ FOCUS AREA	ACTIVITIES
		Lecture & Lab each week, plus:
1	Introduction	Self-evaluation survey (online, no marks)
2	Command-line and Cloud	
3	Fundamental mathematics & statistics	
4	Pairwise sequence alignment	
5	Multiple sequence alignment	
6	Likelihood and Models	**Quiz 1 (20%)**
7	Phylogenetic Inference	
8a	Finding perfect matches: hashing	
Mid Semester Break: April 18 – April 25 inclusive		
8b		**Assignment 1 due April 26th, Friday Week 8b 5pm**
9	Genomics and Sequencing	
10	Read mapping	
11	Genome assembly	
12	RNA-Seq	
13	Extra topics e.g. ChIP-Seq	
exam period		**Assignment 2 due June 14th, 5pm**